

APPLICATION FOR
UNITED STATES LETTERS PATENT
SPECIFICATION

INVENTOR(s): Katsuhito Fujimoto, Atsuko Ohara
and Satoshi Naoi

Title of the Invention: IMAGE PROCESSING APPARATUS AND METHOD
FOR BINARIZING A MUTILEVEL IMAGE

IMAGE PROCESSING APPARATUS AND METHOD FOR BINARIZING A MULTILEVEL IMAGE

Background of the Invention

5 Field of the Invention

The present invention relates to a multilevel image processing technology.

Description of the Related Art

10 Recently, a slip recognition technology using a non-contact type image input device, such as an over-head reader (OHR), has become a key for winning financial OCR (optical character reader) business.

15 An OHR is a stand type image input device provided with a line or area CCD (charge couple diode) as an image element, as shown in Fig. 1A. Compared with a conventional contact type image input device, such as an image scanner, etc., by using an OHR, entry to
20 a slip can be made possible while a user is inputting an image and an image can be inputted while viewing a list of slips. Therefore, work can be performed comfortably.

Compared with an image obtained by a scanner (hereinafter called "a scanner image"), an
25 image obtained by the OHR (hereinafter called "an

00559563-092600

OHR image") suffers from degradation, such as that caused by uneven gradation, reflection, image distortion, etc.

Fig. 1B shows an example of a scanner image, and Fig. 1C shows an example of an OHR image. The OHR image shown in Fig. 1C does not include the reflections of desks, walls, human beings, etc., and it is of fairly good quality for an OHR image. However, compared with the scanner image shown in Fig. 1B, the OHR image has a large degree of uneven gradation and character lines that are more blurred. If an OHR is used, there is also a case where an OHR image with reflections, as shown in Fig. 1D must be handled, since there is a possibility that the reflections of desks, walls, human beings, etc., may be included in an image. The OHR image shown in Fig. 1D is blurred from the right to the left of the image due to reflections and as if the image were gradated. If an OHR is used, the development of a base technology for overcoming such image degradation becomes a major problem.

In order to configure a high-precision binarizing system for an OHR image, it is necessary to obtain a character outline which is resistant against reflection and uneven gradation. Therefore,

constant threshold value binarization is not sufficient and Niblack's local binarization (see Reference 1: IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No. 12, p.1191-1202, 5 1995), etc., must be introduced.

Niblack's local binarization is a system of performing binarization for each pixel assuming that the threshold value of each pixel $T = E + K\sigma$ (E: average gray level of pixels in the vicinity of 10 a target pixel, σ : standard deviation of gray level of the pixels in the vicinity of the target pixel, K: prescribed constant). A rectangular area of $N \times N$ (N is a constant) with the target pixel located at the center is used as the vicinity of the target pixel.

15 However, if a conventional system, such as Niblack's binarization, etc., is used without modification, a black-white flickering noise occurs since all pixels in the vicinity of the pixel have an even gray level inside a background or a thick 20 line.

Fig. 1F shows a binary image obtained by performing Niblack's local binarization ($N=7$, $K=-0.1$) for the OHR image shown in Fig. 1E. According to the conventional binarization system, a black- 25 white flickering noise occurs, as shown in Fig. 1F. Such

Although a method of eliminating the black-white flickering noise which occurs in the case where Niblack's local binarization is applied is described in the previous reference, the method is complex, the process requires many steps and the calculation cost is high, which is a problem.

It is an object of the present invention to eliminate a black-white flickering noise, which is the problem of the conventional binarization system, at a cost that is as low as possible.

One aspect of the present invention comprises a background judgment device and a local binarization device. On receipt of a multilevel image, 25 the background judgment device judges for each

5

10

25

and a stroke separation device. On receipt of a multilevel image, the background judgment device judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the local binarization device locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image. If the ratio of the black pixels in the shape-fixed line element mask, including a target pixel of the obtained binary image is a prescribed value or more, the line element restriction device leaves the black pixels as black pixels or converts all pixels in the line element mask into black pixels. Then, the stroke separation device applies binarization to the partial pattern in the gray scale image, corresponding to the black pixel joint element of the obtained binary image, and divides the pattern into strokes of different gray levels.

Since in any of the configurations described above, the background judgment device roughly judges whether a target pixel is a background pixel prior to the local binarization, the occurrence of a black-white flickering noise can be suppressed at low calculation cost.

Brief Descriptions of the Drawings

Fig. 1A shows the appearance of an over-head reader (OHR).

Fig. 1B shows an example of a scanner image.

5 Fig. 1C shows an example of an OHR image without a shadow.

Fig. 1D shows an example of an OHR image with a shadow.

Fig. 1E shows an example of an OHR image.

10 Fig. 1F shows a binary image obtained by applying Niblack's local binarization to the OHR image shown in Fig. 1E.

Fig. 2A shows the configuration of an image processing apparatus in the first embodiment of
15 the present invention.

Fig. 2B shows the configuration of the image processing apparatus in the second embodiment of the present invention.

Fig. 3 shows an area in the vicinity of a
20 target pixel.

Fig. 4 shows a binary image obtained by applying the process of the second embodiment of the present invention to the OHR image shown in Fig. 1E.

Fig. 5 shows the configuration of the
25 image processing apparatus in the third embodiment of

00569663-092600

the present invention.

Fig. 6 shows an average gray level difference.

Fig. 7 shows a binary image obtained by applying the process of the second embodiment of the present invention to the OHR image shown in Fig. 1D.

Fig. 8 shows a binary image obtained by applying the process of the third embodiment of the present invention to the OHR image shown in Fig. 1D.

Fig. 9 shows the configuration of the image processing apparatus in the fourth embodiment of the present invention.

Fig. 10 shows "r", which is used in the fourth embodiment of the present invention.

Fig. 11 shows the relationship between "r", which is used in the fourth embodiment of the present invention, and a black pixel ratio t.

Fig. 12 shows the configuration of the image processing apparatus in the fifth embodiment of the present invention.

Fig. 13 shows another OHR image.

Fig. 14 shows a binary image obtained by locally binarizing the OHR image shown in Fig. 13.

Fig. 15 shows a binary image obtained by applying the process of the fifth embodiment of the present invention to the OHR image shown in Fig. 13.

5

Fig. 18 shows another OHR image.

10

15

Fig. 22 shows the configuration of the image processing apparatus in the seventh embodiment of the present invention.

20

Fig. 24 shows an information processing device, which is used to configure the image processing apparatus of the present invention.

25

Description of the Preferred Embodiments

The embodiments of the present invention are described with reference to the drawings.

Fig. 2A shows the configuration of the image processing apparatus in the first embodiment of the present invention. This apparatus comprises a background judgment device 101 and a local binarization device 102. On receipt of a multilevel image, the background judgment device 101 judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the local binarization device 102 locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image.

Fig. 2B shows the configuration of the image processing apparatus in the second embodiment of the present invention. This apparatus comprises a background device by standard deviation 201 and a local binarization device 202. On receipt of a multilevel image, the background judgment device by standard deviation 201 judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the local binarization device 102 locally binarizes the

pixel, judges whether the pixel belongs to a background pixel or stroke and outputs a binary image.

The background judgment device by standard deviation 201 judges whether a target pixel is a background pixel, using the standard deviation σ of the gray level of pixels in the vicinity area of the target pixel. Specifically, if $\sigma < \sigma_{\min}$ (predetermined constant), it is considered that the gray level in the vicinity of the target pixel is even and it is judged that the pixel is a background pixel if this condition is met. The vicinity area of a target pixel is the vicinity of the pixel which is a joint area including the target pixel. An example of this is shown in Fig. 3. In Fig. 3, a rectangular area of $N \times N$ with a target pixel 301 located at the center ($N = 7$ in Fig. 3) is the pixel vicinity area 302. The shape of the vicinity area is not limited to a rectangle and a circle, diamond, etc., are also available. It is permissible if the target 301 is not located at the center of the vicinity.

The standard deviation of the gray level of the pixels in the vicinity area 302 is assigned as the standard deviation of the target pixel 301.

The local binarization device 202 performs Niblack's local binarization for only the pixels which the background judgment device 201 judges

not to be a background pixel. Niblack's binarization is a binarization method using an amount which is calculated by $T = E + K\sigma$ using both the average of the gray level E and the standard deviation σ in the vicinity area of a target pixel.

Fig. 1F shows a result obtained by applying Niblack's local binarization to the OHR image shown in Fig. 1E, and Fig. 4 shows a result obtained by applying the process of the image processing apparatus in the second embodiment of the present invention to the image. If the results of Fig. 1F and Fig. 4 are compared, it is found from the process result of the second embodiment, in which Niblack's local binarization is performed after the background judgment by standard deviation, that most of the black-white flickering noises are eliminated. In the process of obtaining the result shown in Fig. 4, a rectangular area of 7×7 with a target pixel located at the center is used as a vicinity area, and it is assumed that $\sigma_{\min} = 10$ and $K = -0.1$.

Fig. 5 shows the configuration of the image processing apparatus in the third embodiment of the present invention. This apparatus comprises a background judgment device by average gray level difference 501 and a local binarization device

5

10

15

Δg = Average gray level of white pixels in the vicinity area - Average gray level of black pixels in the vicinity area.

20

25

10

15

25

the second embodiment of the present invention, which comprises a background judgment device by standard deviation, to the OHR image shown in Fig. 1D, and Fig. 8 shows a result obtained by applying the process of the image processing apparatus in the third
5 embodiment of the present invention, which comprises a background judgment device by average gray level difference, to the OHR image shown in Fig. 1D. In the processes of obtaining the results shown in Figs. 7 and
10 8, a rectangular area of 7×7 with a target pixel located at the center is used as the vicinity area and it is assumed that $\sigma_{\min} = 10$, $K = -0.1$ and $\Delta g_{\min} = 8$.

Since Fig. 1D shows an OHR image with a shadow, there is a possibility that a background may also be judged as a black pixel in a shadowed part, and the gray level difference between a stroke and a background becomes small. Therefore, as shown in Fig. 7, strokes cannot be sufficiently extracted and are blurred in the image processing apparatus,
15 which comprises a background judgment device by standard deviation, while as shown in Fig. 8, strokes can be extracted without blur and a good binarization result with little noise can be obtained in the image processing apparatus, which comprises a
20 background judgment device by average gray level
25

difference.

Although the background judgment device by average gray level difference 501 in the third embodiment of the present invention judges that a target
5 pixel is a background pixel if average gray level difference $\Delta g < \Delta g_{\min}$ (predetermined constant), the average gray level difference also becomes small if there are only stroke areas in the vicinity area. Therefore, a pixel which should naturally be recognized
10 to be a stroke is sometimes judged as being a background pixel. For example, although the characters of "静岡 10" shown in Fig. 1D, which is an original image, are black, the characters of "静岡 10" shown in Fig. 8, which is the process result in the third embodiment of
15 the present invention, are white-punched. In order to prevent this, a process of further judgment about whether the target is a black or white pixel, based on the gray level of the target pixel is added after the judgment by the background judgment device by average
20 gray level difference 501 in the third embodiment of the present invention about whether the target pixel is a background pixel as described earlier. In this way, strokes can be prevented from being white-punched.

Fig. 9 shows the configuration of the
25 image processing apparatus in the fourth embodiment of

the present invention. This apparatus comprises a background judgment device by "r" 901 and a local binarization device 902. On receipt of a multilevel image, the background judgment device by "r" 901 judges for each pixel whether the pixel is a background pixel, and if the pixel is not a background pixel, the local binarization device 902 locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image.

If it is assumed that the standard deviation of the gray level of pixels in the vicinity of a target pixel and the gray level difference in the even area of the target pixel are σ and Δg , respectively, "r" in the background judgment device by "r" 901 is calculated by the following equation.

$$r = \sigma / \Delta g$$

The background judgment device by "r" 901 judges whether the target pixel is a background pixel, using the calculated "r".

Here, "r" is described with reference to Figs. 10 and 11. Fig. 10 is an example of a specific multilevel image and shows a target pixel 1001, a vicinity area 1002, a black area in the vicinity area 1003, a white area in the vicinity area 1004, a stroke

1005 and a background 1006.

It is assumed that the average gray level of the black area in the vicinity 1003 and the average gray level of the white area in the vicinity area 1004 are g_1 and g_2 , respectively. Then, the following equation holds true.

$$\text{Standard deviation } \sigma = r \times |g_1 - g_2| = r \Delta g$$

Specifically, $r = \sigma / \Delta g$. If the black pixel ratio t in the vicinity area is assigned, " r " is represented by the following equation (1), and the " r " and black pixel ratio t have the relationship shown in Fig. 11.

$$\begin{aligned} r &= f(t) \\ &= (t(1-t))^{1/2} \\ &= (t-t^2)^{1/2} \\ &= (1/4 - (t-1/2)^2)^{1/2} \end{aligned} \quad (1)$$

Therefore, if " r " is small, the black pixel ratio t can be small. If the black pixel ratio t is small, specifically, a black area is narrow or there is very little black area, and the target pixel can be judged to be a background pixel. Therefore, if $r < r_{\min}$ (predetermined constant) is satisfied, the target pixel is judged to be a background pixel. Since

r is a quadratic function, r and t are not determined one-to-one. Accordingly, there is a possibility that even if $r < r_{min}$ is satisfied, the target pixel is not a background pixel. However, this can be handled by
5 the background judgment device by average gray level difference in the third embodiment of the present invention executing the same process as that executed to prevent strokes from being white-punched.

The local binarization device 902
10 applies Niblack's local binarization to only pixels which the background judgment device of "r" 901 judges not to be a background pixel.

Fig. 12 shows the configuration of the image processing apparatus in the fifth embodiment of
15 the present invention. This apparatus comprises a background judgment device 1201, a local binarization device 1202 and a line element restriction device (No. 1) 1203. The feature of this embodiment is that the line element restriction device (No. 1) 1203
20 eliminates noises which cannot be formed into a line element since a stroke is composed of line elements, which is formed by joining pixels, and the accuracy of background judgment can be improved.

In the image processing apparatus shown in
25 Fig. 12, on receipt of a multilevel image, the

background judgment device 1201 judges for each pixel whether the pixel is a background pixel, and if the pixel is not a background pixel, the local binarization device 1202 locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image. If all pixels in a shape-fixed line element mask including the target pixel in the obtained binary image are black, the line element restriction device 1203 leaves all the pixels in the line element mask as black, and if they are not so, it modifies all the pixels in the line element mask to white pixels.

Here it is assumed that the line element mask used in the line element restriction device (No. 1) 1203 is a rectangular mask with a horizontal/vertical ratio of 1×3 or 3×1 . Since generally a stroke is composed of three dots or more, a mask of 1×3 or 3×1 can be used and noises which do not amount to the thickness of a stroke can be eliminated. In this way, a stroke on a checkered pattern where one dot of a black pixel and one dot of a white pixel are alternately arrayed can be extracted. The checkered pattern where one dot of a black pixel and one dot of a white pixel are alternately arrayed is recognized as a gray-painted area by human eyes and this

is a technology which an image processing apparatus which receives slips as input images requires.

Fig. 13 shows the OHR image of a character "00230" in a slip, which is written on a checkered pattern which is recognized as a gray-painted area. Fig. 14 shows a result obtained by applying a local binarization to the OHR image shown in Fig. 13, and Fig. 15 shows a result obtained by applying the process of the image processing apparatus in the fifth embodiment of the present invention to the OHR image shown in Fig. 13. In Fig. 14, there are black-white flickering noises in a background area, while in Fig. 15 there are very little black-white flickering noise and clear strokes are extracted.

Fig. 16 shows the configuration of the image processing apparatus in the sixth embodiment of the present invention. This apparatus comprises a background judgment device 1601, a local binarization device 1602 and a line element restriction device (No. 2) 1603. Although the configuration of this embodiment (Fig. 16) is almost the same as that of the fifth embodiment (Fig. 12), they differ in the detailed processes of the line element restriction devices.

In the image processing apparatus shown in Fig. 16,

on receipt of a multilevel image, the background judgment device 1601 judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the

5 local binarization device 1602 locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image. If the ratio of the black pixels in the shape-fixed line element mask including the target pixel in the obtained

10 binary image is a prescribed value or more, the black pixels are left as black pixels or all pixels in the line element mask are converted into black pixels. If the ratio of the black pixels in the line element mask is less than the prescribed value, all pixels in the

15 line element mask are converted into white pixels.

For the line element masks used in the line element restriction device (No. 2) 1603, there are a total of six kinds of masks: four kinds of rectangular masks with horizontal/vertical ratios of 13

20 $\times 1$, 1×13 , 5×3 and 3×5 , with a target pixel located at the center, and two kinds of slant masks of 5×3 and 3×5 . Fig. 17 shows a slant mask of 5×3 1701. The line element restriction device (No. 2) 1603 converts all pixels in the line element mask if 11 pixels

25 are black pixels in the line element mask, and,

009260-096960

Although the image processing apparatus
5 is configured in such a way that the background
judgment device 1601 judges, the local binarization
device 1602 performs a local binarization and the line
element restriction device 1603 eliminates black-
white flickering noises from the obtained binary image,
10 it can also be configured in such a way that the
local binarization device 1602 performs a local
binarization without the background judgment by the
background judgment device 1601 and the line element
restriction device 1603 eliminates the black-white
15 flickering noises from the obtained binary image.

Fig. 18 shows an example of an OHR image with a shadow. The ruled lines in the lower right section of Fig. 18 are blurred. Fig. 19 shows a binary image obtained by applying the process of the third embodiment of the present invention, in which background judgment by average gray level difference is made and a local binarization is performed, to the image shown in Fig. 18. In this case, strokes composing the ruled lines in the lower right section are blurred. Fig. 20 shows a binary image

obtained by applying Niblack's local binarization to the image shown in Fig. 18 without any background judgment. In this case, although there are black-white flickering noises in the background area of Fig. 20, it must be noted that strokes composing the ruled lines are clearly extracted. Furthermore, Fig. 21 shows a binary image obtained by applying the process of the sixth embodiment of the present invention to the image shown in Fig. 18. Although Fig. 21 includes noises, strokes composing the ruled lines are clearly extracted. There is a possibility that good ruled lines may be extracted by eliminating lines short of a prescribed length based on the length restriction of ruled lines, etc., in a subsequent stage.

Fig. 22 shows the configuration of the image processing apparatus in the seventh embodiment of the present invention. This apparatus comprises a background judgment device 2201, a local binarization device 2202, a line element restriction device 2203 and a stroke separation device 2204. The feature of this embodiment is that the stroke separation device 2204 separates two strokes of different gray levels if they touch. According to this embodiment, characters can be accurately extracted from a gray scale image in which a ruled line and a

On receipt of a multilevel image, the background judgment device 2201 judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the local binarization device 2202 locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image. If the ratio of the black pixels in the shape-fixed line element mask including a target pixel of the obtained binary image is a prescribed value or more, the line element restriction device 2203 leaves the black pixels as black pixels or converts all pixels in the line element mask into black pixels. Then, the stroke separation device 2204 calculates the black pixel joint elements of the binary image obtained by the line element restriction device and applies Otsu's binarization (see Reference 2: Technical Report of The Institute of Electronics, Information and Communication Engineers '80/4, Vol. J63-D, No. 4, p.349-356, 1980) to a partial pattern in a gray scale image corresponding to each joint element. If an inter-class dispersion is a prescribed value or more or a dispersion ratio (intra-class dispersion/inter-class dispersion) is less than a

prescribed value, the partial pattern is divided into two strokes.

Fig. 23 shows the concept of the process of the stroke separation device 2204. Otsu's binarization
5 is applied to the partial pattern 2301 obtained by the line element restriction device 2203. Since the partial pattern 2301 is composed of two strokes of different gray levels (a character stroke 2303 and a ruled line stroke 2302), the inter-class
10 dispersion becomes a fairly large value. If the calculated inter-class dispersion is a prescribed value or more, the partial pattern 2301 is divided into two strokes. If the partial pattern is composed of strokes of a small gray level difference, which is not
15 shown in Fig. 23, the stroke separation device 2204 judges that they are the same kind of stroke and the pattern is not divided.

If the stroke separation device 2204 uses a dispersion ratio instead of the inter-
20 class dispersion, it is configured to divide the partial pattern 2301 into two strokes when the dispersion ratio is less than a predetermined value.

The image processing apparatuses described above can be configured using the information
25 processing device (computer) shown in Fig. 24. The

00569563-092600

5

10

20

25

program and data described above in this external storage device and can use the program and data by loading them into the memory 2402, if required.

The medium driving device 2406 drives the
5 portable storage medium 2409. For the portable storage medium 2409, an arbitrary computer-readable storage medium, such as a memory card, a floppy disk, a CD-ROM (compact disk read-only memory), an optical disk, a magneto-optical disk, etc., is used. The program
10 and data described above are stored in this portable storage medium 2409 and can be used by loading them into the memory 2402, if required.

The network connection device 2407
communicates with an external device via an arbitrary
15 network (line), such as a LAN (local area network), etc., and performs a data conversion accompanying communications. The information processing device can receive the program and data described above from the external device via the network
20 connection device 2407 and can use the program and data by loading them into the memory 2402, if required. Although Fig.24 shows a single information processing device, the system can also be implemented by a processing device consisting of a plurality of
25 computers or a plurality of processing devices via a

009260-092600

network.

Fig. 25 shows how to provide a software program, etc., to be run by the information processing device of the present invention. For example, such a
5 program, etc., can be provided by one of the following three methods.

(a) Such a program, etc., is installed and provided in the information processing device, such as a computer, etc. In this case, such a
10 program, etc., are, for example, pre-installed at a plant before shipment.

(b) Such a program, etc., is stored and provided in the portable storage medium 2502. In this case, the program, etc., stored in the portable storage
15 medium 2502 is installed in the external storage device 2405 of the information processing device 2501, such as a computer, etc.

(c) Such a program, etc., is provided from a server in the network 2503. In this case,
20 the information processing device 2501, such as a computer, etc., usually obtains the program, etc., by downloading the program, etc., stored in the server 2504.

In this case, the server 2504 generates a
25 signal for transmitting a program, etc., and transmits

0056953-092500

In this way, according to the present invention, if image data, such as a slip, etc., which are obtained from an input device, such as an OHR, black- white flickering noises, which are conventionally generated in a background area, can be eliminated, and thereby high-accuracy binarization can be realized.

Another aspect of the present invention
25 comprises a background judgment device, a local

binarization device and a line element restriction device. On receipt of a multilevel image, the background judgment device judges for each pixel whether the pixel is a background pixel. If it is judged that the pixel is not a background pixel, the local binarization device locally binarizes the pixel, judges whether the pixel belongs to a background or a stroke and outputs a binary image. If the ratio of black pixels in the shape-fixed line element mask including a target pixel in the obtained binary image is a prescribed value or more, the line element restriction device leaves the black pixels as black pixels or converts all pixels in the line element mask into black pixels. In this way, the line element restriction device can eliminate noises short of a line element and the accuracy of background judgment can be improved.

Another aspect of the present invention comprises a background judgment device, a local binarization device, a line element restriction device and a stroke separation device. On receipt of a multilevel image, the background judgment device judges for each pixel whether the pixel is a background pixel. If the pixel is not a background pixel, the local binarization device performs a local binarization, judges whether the pixel belongs to a background or a

stroke and outputs a binary image. If it is judged that the ratio of the black pixels in the shape-fixed line element mask including a target pixel in the obtained binary image is a prescribed value or more, the line
5 element restriction device leaves the black pixels as black pixels or converts all pixels in the line element mask into black pixels. Then, the stroke separation device applies binarization to the partial pattern in a gray scale image corresponding to the black pixel
10 joint element and divides the partial pattern into strokes of different gray levels. In this way, the stroke separation device can extract high-accuracy characters from a gray scale image in which a plurality of strokes, such as a ruled line, character, etc.,
15 touch.

009260-09600